REPORT RESUMES

ED 010 843 24

A STUDY OF SELECTED SPELLING-TO-SOUND CORRESPONDENCE PATTERNS.

BY- VENEZKY, RICHARD L. WEIR, RUTH H.

STANFORD UNIV., CALIF.

REPORT NUMBER CRP-3090 PUB DATE 66
REPORT NUMBER BR-5-0454
EDRS PRICE MF-$0.18 HC-$3.80 95P.

DESCRIPTORS- *LINGUISTIC PATTERNS, ORTHOGRAPHIC SYMBOLS, *READING RESEARCH, READING PROCESSES, *SPELLING, *PRONUNCIATION, PHONEMES, GRAPHEMES, *MODELS, STANFORD

A LINGUISTIC MODEL WAS DEVELOPED FOR RELATING SPELLING TO SOUND AND TO EXPLORE THOSE FACETS OF ENGLISH ORTHOGRAPHY WHICH MIGHT RELATE TO THE READING PROCESS. A DETAILED ANALYSIS OF THE BASIS OF THE ORTHOGRAPHY WAS MADE WHICH INCLUDED DISCUSSIONS OF THE GRAPHEME-PHONEME PARALLEL, RELATIONAL UNITS, MARKERS, AND GRAPHEMIC ALTERNATIONS. REVISIONS AND EXTENSIONS WERE MADE TO FORMULATE THIS IDEAL SYSTEM FOR TRANSLATING FROM SPELLING TO SOUND, BASED ON AN ORIGINAL GROUP OF 20,000 WORDS. TENTATIVE IMPLICATIONS POINTED TO THE POSSIBILITY OF TEACHING VARIOUS PRONUNCIATIONS NOT BY THE SIMPLE-SEQUENCE METHOD WHERE VARIANT PRONUNCIATIONS ARE PRESENTED SEQUENTIALLY, BUT BY OFFERING ALL THE DIFFERENT PRONUNCIATIONS AT ONCE, WORKING WITH PAIRS OF WORDS WHICH SHOW THE DIFFERENT PRONUNCIATIONS. IT WAS FELT THAT THE POTENTIAL GENERALIZATION DERIVED FROM THIS DIFFERENTIATION APPROACH SHOULD EFFECT BETTER RESULTS THAN THE SIMPLE-SEQUENCE METHOD. (GD)

A STUDY OF SELECTED SPELLING-TO-SOUND CORRESPONDENCE PATTERNS

Cooperative Research Project No. 3090

Richard L. Venezky
Ruth H. Weir

Stanford University
Stanford, California

1966

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

The research reported herein was supported by the Cooperative Research Program of the Office of Education, U. S. Department of Health, Education, and Welfare.

A STUDY OF SELECTED SPELLING-TO-SOUND
CORRESPONDENCE PATTERNS

Cooperative Research Project No. 3090

Richard L. Venezky
Ruth H. Weir

The research reported herein was supported by the Cooperative Research Program of the Office of Education, U. S. Department of Health, Education, and Welfare.

Stanford University
Stanford, California

1966

PREFACE

The material presented here is dedicated to the memory of Ruth H. Weir whose tragic and untimely passing in November of 1965 deprived me of a friend and colleague,and the academic world of an esteemed scholar. Since Mrs. Weir died before this report was prepared, the opinions expressed herein probably do not fully represent her perceptive and challenging [illegible] report is more than a formality to meet contractual obligations to the Office of Education. This project is based upon two previous research efforts by the same authors (under support from the U. S. Office of Education) and upon my doctoral dissertation research which Mrs. Weir supervised. While the verbiage in this report is entirely my own, many of the ideas were hers. Memories of people fade with time, but good ideas are self perpetuating. It is in this sense that I dedicate this present work to her memory.

Richard L. Venezky

CONTENTS

Problem on which the Research was Focused 1

Objectives and Hypothesis.........................1

Related Research..................................4

Procedure...5

Results

1. The Basis of English Orthography........6
2. Consonant Correspondences and Alternations - A Summary......................34
3. Vowel Correspondences and Alternations - A Summary.............................54
4. The Morphophonemics of Voiced-Voiceless Alternations in American English Fricatives.............................70

Conclusions and Suggestions for Further Research...81

List of References...............................86

Problem on Which the Research Was Focused

The research reported here is part of a continuing investigation of English orthography and its relationship to sound. Our emphasis has been upon the structure of spelling-to-sound correspondences and the various elements - graphemic, morphemic, and phonemic - which enter into this system. In previous reports we presented a model for spelling-to-sound relationships and described the application of this model to certain orthographic patterns. The graphemic basis of the orthography was discussed in general terms, along with the patterning of vowel spellings and the influence of /w/, /l/, and /r/ upon vowel correspondences. In addition, an exhaustive analysis of the final e pattern, a preliminary analysis of consonant spelling-to-morphophoneme patterns, and a number of morphophonemic patterns were presented. In this report we present a detailed analysis of the basis of English orthography, a revised and extended description of consonant and vowel spelling-to-morphophoneme patterns, consonant morphophonemic alternations, and a discussion of the application of the results of this work to the teaching of reading.

Objectives and Hypothesis

We have attempted in this work not only to develop a model for relating spelling-to-sound, but to explore any facet of the orthography that might relate to the reading process. While the improvement of the teaching of reading is one of our primary goals, we have attempted to separate this in our research from the analysis of the orthographic system. Our strategy has been to analyze the orthography and its relationship to sound as an independent system without slanting the investigation towards any particular discipline. From the results of this analysis, however, we have attempted to abstract those patterns which either can be related directly to the teaching of reading or can be used to investigate reading habits.

Early in this work we advanced the hypothesis that English spelling is not simply a defective phonemic system for transcribing speech, but instead a more complex and more regular set of patterns in which both phonemic and morphemic elements share leading roles. Such an idea, while not held universally by modern linguists, nevertheless does not originate from this work. Henry Bradley, one of the editors of the Oxford English Dictionary, described English orthography in similar terms as early as 1914.[1] More recently, C. F. Hockett has written

> The complexities of English spelling cannot be accounted for completely on the assumption that the system is phonemic with irregularities of the sort listed . . .[above]. It is necessary to assume that the system is partly phonemic and partly morphemic.[2]

---

1. Henry Bradley, "On the relations between spoken and written language with special reference to English" British Academy 6.211-32 (1914).

2. Charles F. Hockett, A course in modern linguistics (New York, 1956), p. 542.

One consequence of this hypothesis is that attempts to relate spelling directly to sound (or sound directly to spelling) cannot account for much of the patterning in the orthography, that only a view that relates spelling to some intermediate level from which both phonemic and morphemic elements can be derived will productively accommodate the existing system. For this reason we developed a model for mapping first from spelling onto a morphophonemic level, and then from there, through various morphophonemic alternations into sound. This is not to claim that the orthography is a morphophonemic system, but only to say that the orthography fits more snuggly into a morphophonemic model than into a direct spelling-to-sound one.

The second consequence of the hypotheses stated above is that the categories of regular and irregular correspondences are not by themselves meaningful classes for describing the orthography. A correspondence that is (by some definition) irregular from a phonemic standpoint may be regular from a morphemic standpoint. Furthermore, the dichotomy of regular and irregular is based upon arbitrary divisions that ultimately derive from frequency of occurrence. Yet some low frequency correspondences, though classified as irregular, may still represent describable patterns.

The notions of regular and irregular compose, for the most part, an arbitrary partitioning imposed upon a set of relationships. Within the set of relationships, the actual patterns are invariant, since they describe what factors, if any, can be used to predict sound from spelling. Which of these patterns will be classed as regular and which as irregular, however, is variable, since the size and nature of the corpus and the ultimate goal of the classification affect the basis for selection. If for example, the corpus upon which the correspondences are based is limited to the most common 5,000 words in English texts, more invariant patterns will appear than will if the most common 20,000 words were used. Therefore, the classes "regular" and "irregular" must always be qualified by the corpus employed and the principle for dichotomizing.

Related Research

Most of the material presented here is based upon R. L. Venezky's dissertation, "A study of English spelling to sound correspondences on historical principles," Stanford University, 1965, and upon two previous research projects by the two investigators (Weir, 1964; Weir and Venezky, 1965). Current American research on reading, although mostly of indirect interest to our work, has been described in the Project Literacy Reports (see Levin, 1963, 1964, 1965). Modern studies of spelling-to-sound correspondences have been reported by Hall (1961, 1962), Soffietti (1955), and Waldo (1964). Notions on the relation between writing and speech can be found in Bradley (1914), Bolinger (1946), Uldall (1944), Pulgram (1951), Vachek (1962), and Stetson (1937). Francis (1958) contains the most complete discussion of English graphemics published to date.

Procedure

The initial data upon which this study is based were obtained by R. Venezky through a computer-aided analysis of the spelling-to-sound correspondences in approximately 20,000 different English words.[1] Output from a program written for the CDC 1604-A included a complete tabulation of spelling-to-sound correspondences for this corpus, along with the word lists for each correspondence. In addition, a reversed spelling listing and a reversed pronunciation listing of the corpus were obtained. Besides the analysis of this material, studies were made of the Hanley Rime Collection, a large unpublished collection of English and American rimes.

The phonology used in this report is based upon the system presented in H. Kurath, *A phonology and prosody of Modern English* (Ann Arbor, 1964). Actual pronunciations for the corpus are derived primarily from the *Thorndike century senior dictionary* (New York, 1941), and Kenyon and Knott, *A pronouncing dictionary of American English* (Springfield, Mass., 1951).

---

[1] See R. L. Venezky, "A computer program for deriving spelling-to-sound correspondences," *A basic research program on reading*, Cornell University, 1963, and R. Weir and R. Venezky, *Rules to aid in the teaching of reading*, Final Report, Cooperative Research Project No. 2584, Stanford University, 1965, Appendix A.

RESULTS - 1

The Basis of English Orthography

Orthography, regardless of the status we assign to it within the science of linguistics, is an important feature of the English language. It is the source of most of our knowledge of Old and Middle English; it is the principle medium through which foreigners come in contact with Present-Day English, and it is an object of frequent scrutiny by linguistics who labor over the problems of teaching spelling and reading. In spite of this position, however, there has been published neither a comprehensive study of Modern English spelling-to-sound correspondences nor a detailed history of English spelling. The available information on these topics must be gathered from a multitude of diverse sources and pieced together much as ancient vases are reconstructed from unsorted potsherds.

The earliest writings on English orthography are based upon an alphabetic principle derived from the fourth and sixth century Roman grammarians. Each letter of the alphabet has, besides its name (nomen) and appearance (figura), a power (potestas) or sound, and a description of the orthography involves simply a classification of the letters according to their powers.[1] Consequently, orthography from the time of

---

[1] For further notes on the description of letters according to nomen, figura, and potestas, see Einar Haugen, The first grammatical treatise: the earliest Germanic philology (Baltimore, 1950), pp. 41-2.

Alfred to the present day has been delimited by the letters and their powers. So ingrained has this principle become that some contemporary linguists have attempted, by substituting grapheme for letter, to sanctify it with the countenance of linguistic science without examining how unsound it is. The outcome of this position has been the view that English orthography is a faulty phonetic system, and consequently much of the graphemic patterning and almost all of the morphemic patterning in the orthography has been ignored.[1]

Some critical exceptions to this view are found in the writings of Francis, Hockett and Vachek.[2,3,4] While all three of these have recognized that the orthography is more complex than an irregular letter-to-sound system, only Francis has attempted to analyze the relation between spelling and sound and to enumerate the non-phonemic elements which enter into this relationship. Most other linguists have paid little

---

1. This view is especially evident in Otto Jespersen, A modern English grammar on historical principles (Heidelberg, 1909-31), I, 3, and Ernst Pulgram, "Phoneme and grapheme: a parallel", Word 7 (1951) 15-20.

2. Winthrop N. Francis, The structure of American English (New York, 1958), 450-69.

3. Charles F. Hockett, A course in modern linguistics. (New York, 1956), 542.

4. Joseph Vachek, "Two chapters on written English," Brno Studies in English I, 8 (Praha, 1959).

attention to English orthography. Bloomfield, for example, held that writing was not a part of language, but simply an imperfect image of speech, and even though he wrote at length on the teaching of reading, he maintained that the orthography was a grossly irregular phonetic system.[1]

Spelling reformers, who have contributed an enormous quantity of literature on English orthography, have seldom analyzed the object of their scorn beyond the more common examples of scribal pedantry. Their arguments were, and are still, based upon the a priori assumption that alphabets should, and by some right ought to be perfectly phonetic. It is no surprise, therefore, that most spelling reformers have concerned themselves with direct letter-to-sound relationships and have ignored all other facets of the writing system.[2]

---

1. "Although our writing is alphabetic, it contains so many deviations from the alphabetic principle as to present a real problem, whose solution has been indefinitely postponed by our educators' ignorance of the relation of writing and speech..." Language (New York, 1933), 500-501

2. "...everyone...has to admit that of all languages of culture English has the most antiquated, inconsistent, and illogical spelling." Robert E. Zachrisson, Anglic, a new agreed simplified English spelling, (Uppsala, 1930), 10. For a slightly less polemic statement of the spelling reformers' stance, see Zachrisson's "Four hundred years of English spelling reform," Studia Neophilologica 4 (1931), 1-69.

In two previous papers we presented preliminary analyses of the basis of English orthography.[1] Included were notes on the grapheme-phoneme parallel and the difference between relational units and markers. In this section we present a complete analysis of the basis of the orthographic system, including discussions of the grapheme-phoneme parallel, relational units, markers, and graphemic alternations. This material forms the groundwork for the results presented in the next two sections, and contains, in some cases, revisions and improvements upon the material which we presented in the two earlier reports.

Selection of Units

The Grapheme-Phoneme Parallel

The description of spelling-to-sound relationships begins with certain units on the spelling or graphic level and ends, for convenience, not with actual speech sounds, but with phonemes. Traditional studies in this area have generally employed the term grapheme as a parallel to phoneme and various attempts have been made in recent years to formalize graphemics or to elaborate on the grapheme-phoneme parallel.[1]

---

[1] See R. H. Weir, "Formulation of grapheme-phoneme correspondence rules to aid in the teaching of reading;" Final report, Cooperative Research project No. S-039 (Stanford, 1964) and R. H. Weir and R. L. Venezky, "Rules to aid in the teaching of reading," Final report, Cooperative Research project No. 2584 (Stanford, 1965).

> Borrowing some ideas and methods from phonemics and morphemics, for instance, we could conclude that each of the various letters has two or more different shapes, which seem to be in complementary distribution or free variation. This in turn suggests that each different shape can be called an *allograph*, and a family of allographs a *grapheme*.... Taking inventory of the segmental graphemes of standard English writing or printing, we find that there are thirty-seven of them, which can be classified in two groups:
>
> (a) Twenty-six letters of the alphabet ⟨abc. . . z⟩
>
> (b) Eleven marks of punctuation ⟨, ; : . ? ! - " () --⟩
>
> In addition we must include space, a sort of zero grapheme.[2]

Even with this description, the graphemic and phonemic systems are far from parallel. While various single graphemes can be mapped (through intermediate levels) into single phonemes, there are also grapheme clusters that operate as single units, like *th* and *ch*. Since we cannot deduce the behavior of these clusters from the behavior of their constituents, we must conclude that there exist units higher than the grapheme on the graphic level, like *ch* and *th*, which correspond more closely to the phoneme than does the grapheme. An adequate description of English graphics should, therefore, describe such units. To meet this need, Hall differentiates between simple and compound graphemes.

---

[1] The earliest mention of the grapheme-phoneme parallel that I can find is in F. de Saussure, *Course in general linguistics*, trans. Wade Baskin (New York, 1959), 23, 30. Saussure saw four features in the letters (he did not use the term *grapheme*) which paralleled features in the phonemic system. Pulgram upped the count of parallel features to nine in 1951 (see page 8 . For a summary of linguistic writings on the grapheme-phoneme parallel, see John C. McLaughlin, *A graphemic-phonemic study of a Middle English manuscript* (The Hague, 1963), 24-28.

[2] W. N. Francis, *Structure*, p. 436.

> The compound graphemes of English include a great many sequences of vowel letters, graphic diphthongs such as *ae*, *ai*, *au*, *ea*, *ei*, *eo*, . . . *eu*, *ie*, *oa*, *oi*, *ou*, *ue*, in addition to double vowel letters like *ee* or *oo*. There are also certain combinations of consonant letters which function as single units and hence must be considered as compound graphemes: e.g., *ch*, *gh*, *ph*, *rh*, *sh*, *th*, *ng*, and again, all the double consonant letters such as *bb*, *dd*, etc.[1]

Even this division, however, does not create a workable parallel between grapheme and phoneme. Certain graphemes function solely as markers, that is, do not themselves enter into correspondences, but rather, mark one of several possible correspondences for other graphemes or grapheme clusters. Final *e*, for example, performs at least six different marking functions (see pages 23 and 24). It indicates vowel correspondences as in *mate* and *cute* (cf. *mat* and *cut*), it marks consonant correspondences as in *trace*, *change*, and *bathe* (cf. *bath*), and preserves graphotactical patterns as in *have*, *tow*, and *glue*. The phonemic level has no parallel to this.

---

[1] Robert A. Hall, Jr., *Sound and spelling in English* (Philadelphia, 1961), 14.

One other difference between graphemes and phonemes is that phonemes are language dependent, functionally-defined units, while graphemes are not necessarily language dependent nor functionally defined. Almost all speech communities which use the Roman alphabet have the same graphemic system, according to current definitions of graphemics, yet their phonemic systems are vastly different. Furthermore, a new grapheme can be added at will to the graphemic system, regardless of whether or not it contrasts functionally with existing graphemes. The Anglo-Norman scribes, for example, re-introduced q into English orthography. Yet this grapheme performed the same function as k which remained in the English writing system.[1] The parallel situation does not exist on the phonemic level; in fact, the opposite is true. A new phoneme comes into existence only when a new contrast appears.

The term grapheme refers to letter classes. The individual letters, as written, typed, or in any other way produced, are called graphs, and each graph is classed as one of the twenty-six graphemes in the current alphabet. The decision to label two graphs as same or different depends in no way upon whether they function as the same or different form. /klæn/, for example, can be spelled *clan* or *klan*, yet *k* and *c* even in this environment are not classed as members of the same grapheme. The difference between these two graphs, therefore, cannot depend upon the relationship of spelling to sound or upon any other language dependent feature, but only upon the graphic features of the two forms. This process of classification is vastly different from the process by which two phones are classed as same or different. Form alone is considering in classing graphs into grapheme patterns, but both form and function are essential for classing phones as members of phoneme classes.[1]

---

[1] Latin q occurred before v as a spelling for /kw/ in some early Old English manuscripts (see OED, s.v. q), but cw was a more common spelling for OE /kw/. After the Conquest q was reintroduced in English orthography by Anglo-Norman scribes and by the end of the thirteenth century qu had replaced cw and kw as a spelling for /kw/.

Type of Graphemic Units

In the remainder of this report *grapheme* will refer to one of the alphabetic characters *a* through *z* and the term *graphemic* level will be used as a general reference for the spelling level.[1] The actual language-dependent units on the graphemic level which are significant for the prediction of sound are called functional units, and are divided into two classes: relational units and markers.

---

[1] The discrimination of graphemes has been studied by a number of psychologists in recent years and attempts have even been made to formulate distinctive features for letters. See especially Eleanor J. Gibson, *et al.*, "An analysis of critical features of letters, tested by a confusion matrix," Harry Levin, ed. *A basic research program on reading* (Cornell, 1963).

> A relational unit is a string of one or more graphemes which has a morphophonemic correspondent which cannot be predicted from the behavior of the unit's smaller graphemic components.
>
> A marker is a cluster of one or more graphemes whose primary function is to indicate the correspondences of relational units or to preserve a graphotactical or morphological pattern. It is mapped into a zero morphophoneme.

---

[1] Grapheme is not an ideal term either for the discussion of spelling-to-sound correspondences, but I find the only plausible substitute, letter, more objectionable. The case against letter was summarized by Angus McIntosh, "The analysis of written Middle English," Trans. Phil. Soc., 1956, p. 43. "The word 'letter' in everyday use is ambiguous. It may be employed in the sense of 'grapheme' as when we say that the word swilk begins with 'the letter s'; at other times it is used of the particular allographic form of a grapheme may have in a given context; e.g., when we speak of 'the s used in final position in Greek'; again, it may be used of a single instance of an allograph, as when we say 'that's a badly formed letter'. . . ."

Relational Units

The same grapheme or grapheme cluster can be part of more than one relational unit. Thus, gn in cognac and poignant is a single relational unit which corresponds to the morphophonemic cluster :nj:, but gn in sign and malign is not a relational unit, but rather a combination of two relational units which correspond to the morphophonemes :g: and :n:. Morphophonemic rules map :g: into either :ø: or into :g:, depending upon allomorphic considerations (cf. signal and malignant). The selection of relational units is based upon function and composition. Any string of graphemes that corresponds to a non-zero morphophoneme is a potential relational unit. However, only those strings whose morphophonemic correspondences cannot be predicted by general rules based upon smaller units contained in the string are classed as relational units. ch in chair, for example, is a relational unit since the morphophoneme :č: cannot be predicted from general rules based upon c and h separately. Geminate consonant clusters, however, are not single relational units since their morphophonemic forms can be obtained from rules based upon their separate constituents. (The leveling of clusters like :ff: to :f: can be accounted for by a general phonotactical rule.)

Consonant relational units, furthermore, are classed as functionally simple or functionally compound.[1] This distinction is needed for an accurate statement of a number of rules. In the sequence *vowel* + *consonant* + *final e*, for example, *vowel* is generally mapped into its free alternate if *consonant* is a functionally simple unit (or this type of unit plus *l* or *r*), and into its checked alternate if *consonant* is a functionally compound unit, or a cluster.[1] Thus,

---

[1] The functionally compound units are tch, dg, ck, and x. The remaining consonant units (including sh, th, ch, etc.) are functionally simple.

Free alternate	*Checked alternate*
b*a*ke	*a*xe
*a*che	b*a*dge
conc*e*de	*e*dge
cl*o*the	h*o*dge

Vowel units are classed as primary (*a*, *e*, *i*, *y*, *o*, *u*) or secondary (all others).

Examples of markers are the final *e* in *mate* and *peace*, the *u* in *guest* and the *k* in *trafficking*, all of which indicate the pronunciation of a preceding grapheme. *ue* in *plague*, *catalogue*, etc., is a sequence of two markers. The *u* after *g* marks the correspondence *g*→:g: rather than *g*→:ǰ:. Since *u* does not appear in final position in English words (except after a vowel in a few recent borrowings) a final *e* is added, as in *continue* and *blue* (cf. the alternations ou/ow, au/aw, pages 31 and 32). Only graphemes mapped into zero can be classed as markers (this is a necessary, but not sufficient condition). Graphemes with non-zero morphophonemic correspondences, however, though properly classed as relational units, can also perform marking functions. The *i* in *city*, for example, besides corresponding to :I:, marks the correspondence *c*→:s:. A geminate consonant cluster also performs a marking function since it regularly indicates the correspondence of the preceding vowel.

---

[1] Free and checked vowels are described in Hans Kurath, *A phonology and prosody of Modern English* (Ann Arbor, 1964), 17-20. A description of vowel mappings is given in Weir and Venezky, 1965, pp. 21-28.

The strongest evidence for a separate class of markers in English orthography is found in orthographic alternation patterns. For example, final *e* as a marker for the pronunciation of a preceding *c* or *g* is dropped before a suffix which begins with a letter that will perform the same function as the *e*. Therefore, *notice* drops the final *e* before *ing* (*noticing*) since *i* also marks the correspondence *c*→:s:, but retains the *e* before *able* since *noticable* would have *c*→:k:. Similarly, the *e* added to an otherwise terminal *u* is dropped before any suffix since the only functions of the *e* is to avoid having word-final *u*, e.g., *argue*, *arguing*.

Relational units are classed as consonants or vowels depending upon the class of the morphophonemes into which they are mapped (glides are classed as consonants). Some *relational units* are classed as both consonant and vowel, e.g., *u* in *language* (consonant) and *during* (vowel). Within these classes major and minor patterns are distinguished on the basis of frequency of occurrence. Thus, *ch* is classed as a major consonant unit, but *kh* (khaki) is classed as a minor unit. While the major-minor classification may appear arbitrary, it distinguishes frequently occurring, productive patterns from infrequent patterns which generally occur only in a small number of borrowings.

In the traditional treatment of spelling, forms like *gu* (*guard*), *ti* (*nation*), *di* (*soldier*), and *a . . .e* (*bake*) are adopted as basic spelling units.[1] Although such a course has not been adopted here, it is interesting and informative, nevertheless, to extend this treatment to some of its logical conclusions. In the units just mentioned, so-called silent letters are always included as parts of other spelling units, regardless of the functions which they perform. Therefore, the *b* in *debt* and the final *e* in *peace* would fall into the same basket, in that they are both silent--the units, presumably, being *bt* and *ce*. This treatment, unfortunately, obscures a basic difference between these two silent letters, namely, that the *e* at the end of *peace* marks the correspondence *c*→:s: while the *b* in *debt* is a functionless scribal insertion. The *e* at the end of *peace* is part of the general pattern of markers in English orthography; the *b* in *debt*, on the other hand, is an unproductive, isolated occurrence.

Consider, furthermore, the traditional enumeration of spelling units as reflected in such words as *clothe* and *pace*. Both words have, by the traditional treatment, a silent *e* which must be attached to another grapheme, yet it is not clear which unit either e should be attached to, since in both words it marks two separate patterns. In clothe e marks the correspondences o→:o: and th→:ð:; in pace it marks a→:e: and c→:s:. The traditionalist is faced with a dilemma here: are the units o . . . e/a . . . e or the/ce? Or shall we take a fine razor and split e into two parts so that both alternatives can be taken? Without treating e as a member of a class of markers that are separate from the relational units, no realistic solution is possible.

---

[1] See especially Hall, 1961, pp.17-22.

The last of the major problems inherent in the traditional view is that an unnecessarily large number of basic units are created by the merger of markers with relational units. Besides classing all of the geminate consonant clusters as basic units, the following, among others, must also be recognized as basic:

the	breathe	ce	trace
dge	edge	lle	belle
le	able	gne	cologne
re	acre	ffe	giraffe
ye	bye	aoh	pharaoh
ge	change	ah	pariah
xe	axe	eh	vehement
ve	love		

All of the units above must be classed as basic units by the traditional view, yet their behavior can be predicted from the behavior of their components. The final e in the clusters shown above is a marker whose various functions are explained on page 23. Geminate clusters (belle, giraffe) can be handled by considering each consonant separately and leveling the resulting geminate morphophonemic clusters by a single rule. h in pharaoh and pariah is a marker (see page 23), while in vehement it is deleted by a phonotactical rule (cf. vehicle:vehicular).

Enumeration of Relation Units

The consonant and vowel relational units are shown below. Units followed by an asterisk are discussed in the following paragraphs.

Major Relational Units

Consonants						Vowels			
simple					compound	primary	secondary		
b	gh	n	s	w*	ck*	a	ai/ay	ie	ue
c	h	p	sh	wh	dg	e	au/aw	oa	ui
ch	j	ph	t	y*	tch	i	ea	oe	
d	k	q	th	z	x	o	ee	oi/oy	
f	l	r	u*			u	ei/ey	oo	
g	m	rh	v			y	eu/ew	ou/ow	

Minor Relational Units

Consonants		Vowels
simple	compound	secondary
kh	gn	ae
sch		eau
		eo
		uy

u is a consonant unit when it corresponds to :w: as in quack, language and assuage. It may also be a vowel unit, or part of a vowel unit (ou), or a marker (guest, plague).

w is a consonant unit when it corresponds to :w:, e.g., warm, beware. It also appears as part of a vowel unit (ow, aw) but never as a vowel unit itself.

y is a consonant unit when it corresponds to :j:, e.g., yes, beyond. It also appears as a vowel unit, and as part of a vowel unit, e.g., cycle, boy.

ck is a consonant unit in words like rack and tack. In instances like picnicking, however, the k is a marker. That ck in picnicking is identical to the relational unit ck is immaterial, since the base form picnic ends in c, not ck.

Examples of Minor Units

kh	khaki, khan	ae	aesthetic, algae
sch	schist, schwa, seneschal	eau	bureau, plateau
gn	cognac, poignant	eo	jeopardy, leopard
		uy	buy, guy

Markers

Introduction

The primary functions of markers, as mentioned above, are to indicate the correspondences of other graphemes and to preserve graphotactical or morphological patterns. For example, v does not occur in final position in English; where it would a final e has been added, e.g., have, love. Various relational units also serve as markers, but since they have non-zero morphophonemic correspondents, they are classed as relational units rather than as markers. In the discussion below markers are classed according to the functions they perform.

Markers of Vowel Correspondences

1. In the word-final patterns *VCe* and *VCle*, *e* generally indicates the free pronunciation of *V*.[1] Thus, *mate*:*mat*, *mete*:*met*, *site*:*sit*, *note*:*not*, *cute*:*cut*.

2. *h*. Post vocalic *h* before a consonant or juncture is classed as a marker, even though no clear marking function can be determined from the few existing occurrences of this form: *ah*, *hurrah*, *oh*, *ohm*, *pharaoh*, *pariah*, *hallelujah*. (The alternative is to class *ah*, *oh*, and *aoh* as relational units.) In *prohibition*, *vehement*, etc., *h* is mapped into :h: and then deleted by a general morphophonemic rule.

Markers of Consonant Correspondences

1. *e*. In final position after *c*, *e* marks the correspondence *c*→:s:. In adding suffixes which do not begin with *i*, *y*, or *e*, the *e* is retained as a marker, e.g., *trace*:*traceable*:*tracing*.[2]

2. *e*. In final position after *consonant* + *l* or *consonant* + *r*, *e* marks a syllabic *l* or *r*. Thus, *able*, *table*, *acre*, *massacre*.

---

[1] See Weir, 1964, pp. 42-45.

[2] *e* performs two marking functions in *trace*; it marks the correspondences *a*→:e: and *c*→:s:.

Consonant + re spellings are rare in American English, occurring regularly only in the patterns cre and gre where the more normal cer and ger would indicate the wrong pronunciation for c and g.

3. e. In final position after g, e marks the correspondence g→:ǰ:, e.g., age, college. After dg, e is a graphotactical marker, since dg does not occur in final position. In adding suffixes which do not begin with the i, y, or e, the e is retained as a marker, e.g., change:changeable:changing. In final position after th, e marks the correspondence th→:ð:. Thus, bath:bathe, breath:breathe, teeth:teethe, cloth:clothe, wreath:wreathe.

4. u. After g, u is occasionally used to mark the correspondence g→:g: (cf. the use of e to mark g→:ǰ:, above.) For example, guess, guide, guest, guise, catalogue, guilt, plague, guild.

5. k. A k is inserted between a final c and a suffix beginning with i, y, or e, to mark the correspondence c→:k:. Thus, picnic:picnicking, traffic:trafficking, panic:panicky, shellac:shellacked.[1]

6. Gemination. A geminate consonant cluster regularly marks a checked correspondent for a preceding vowel. Compare,

---

[1] But, irregularly, arc:arcing, disc:discing:disced. In lyricism (cf. lyric, lyrical) the phonetic change in Latin of [k]→[s] obviated the need for the k marker.

for example, *anal*:*annals*, *fetal*:*fetter*, *hypo*:*hippo*, *rotor*:*rotter* (*blackguard*), *super*:*supper*.

Markers of Graphotactical Patterns

1. *e*. After what would otherwise be a final *v* or *u*, an *e* is added. This practice developed during the late Middle English period, due partially to the graphical identity of *u* and *v*. Thus, *love*, *glue*, *have*, *plague*, *glove*, *continue*.

2. *e*. To limit the two letter words in English to a select group of common words, a final *e* has been added after a vowel in what would otherwise have been a two letter word. Thus, *see*, *doe*, *toe*, *rye*, etc. On the addition of *e* after *o*, Webster's 1852 *Dictionary* had the following explanation:

> Woe.--This word takes the final *e*, like *doe*, *foe*, *hoe*, *sloe*, *toe*, and all similar nouns of one syllable. The termination in *o* belongs among monosyllables, to the other parts of speech, as *potato*, *tomato*, etc.[1]

3. *al*. The cluster *al* is inserted between a final adjectival *ic* and the adverbial *ly* if an adjective in *ical* does not exist. Thus, *basic*:*basically*, *rustic*:*rustically*, *scenic*:*scenically*. Generally, the *al* is mapped into zero in such forms.

---

[1] *Webster's American dictionary of the English language*, ed. Chauncey A. Goodrich (2d ed. rev.; Springfield, Mass., 1852), p. 81. Two letter words ending in a consonant were lengthened, not by adding a final *e*, but rather, by doubling the final consonant; thus, *ebb*, *add*, *egg*, *inn*. (These are the only examples of final *bb*, *dd*, *gg*, *nn* in Modern English spelling.)

Spelling pronunciations, however, have introduced forms in which the al is pronounced.

Markers of Morphemic Patterns

1. e. After a number of words which would otherwise end in non-morphemic s, an e has been added to avoid the appearance of a final, inflectional s. Thus, moose, goose, mouse. After the cluster rs, e also marks the correspondence s→:s: in contrast to s→:z:, e.g., hearse:hers. Non-morphemic s is generally avoided in final position in nouns, adjectives (by analogy to nouns), and verbs under the following conditions:

(1) when preceded by a consonant

collapse
eclipse
else
false
dense
sense
hearse
coarse

(2) when preceded by a compound vowel

goose
moose
house
mouse
praise
raise
noise
poise

The only true exceptions to the restrictions are summons (n., sg.), lens, and words ending in ous. Summons, an Anglo-French adoption, had variant spellings in nce and nse through the seventeenth century (see OED, s.v.), but apparently through analogy (both spelling and pronunciation) with the identical form, summons became the standard spelling. Lens is a late seventeenth century Latin adoption for which no variant spellings in *se* were found in the data used for this study. If, however, this word were a phonological rather than a spelling adoption, then *lense* would give the incorrect pronunciation /léns/. (This is also true of *summonse*.) The suffix *ous*, if spelled *ouse* would be pronounced /aus/ as in *mouse* and *house*. Apparent exceptions to (1) and (2) above are words like *hydraulics*, *mathematics*, *mumps*, and *billiards*, which function as both singulars and plurals. In these, however, the *s* can be considered as morphemic.

The only situations where the *e* marker is not added is where *s* is preceded by a simple vowel spelling, but here *e* cannot be added because it would tend to mark an incorrect pronunciation for the simple vowels. Compare *us*:*use*, *his*:*rise*, *locus*:*recluse*, *tennis*:*improvise*.

The introduction of the *e* marker for non-morphemic *s* appears to have begun in the middle of the fifteenth century and to have become standard by the middle of the sixteenth century.

Graphemic Alternations

i and *y*

Vocalic *i* and *y* (which can be considered as members of the same functional unit) occur in complementary distribution: *y* in final position and *i* in initial and medial position. There are, however, numerous exceptions to this pattern. A considerable number of Greek and Latin borrowings retain *y* in medial position, e.g., *embryo*, *asylum*, *rhythm*. In addition, some medial *y*'s have resulted from inorganic final *e*'s, e.g., *rye* and *bye*, and from scribal pedantry, as in *rhyme* which was converted from the earlier *rime* on the mistaken analogy with L. *rhythm*. Initial *y* (vowel) occurs in the forms *ycleped*, *yttrium*, and *ytterbium*, while final *i* occurs in several distinct patterns. One is the plural of Latin borrowings whose singulars end in *us*, e.g., *alumni*, *fungi*, *magi*, and *stimuli* (*alibi* and *quasi* are singular in Latin). Similarly, it occurs as the plural of Italian borrowings whose singulars (used in Italian only) end in *o*, e.g., *broccoli* (It. sg. *broccolo*), confetti (It. sg. *confetto*), *ravioli* (It. sg. *raviolo*), and *spaghetti* (It. sg. *spaghetto*). *Macaroni*, another Italian borrowing, is singular in Italian. Of the remaining words ending in *i*, *alkali*, *charivari*, *houri*, and *potpourri* are French, *anti* is Greek, *ski* is Norwegian, *chilli* is Mexican, *khaki* is Urdu, *mufti* is Arabic, *yogi* is Hindustani, and *taxi* is an early twentieth century American coining. Many of these words had alternate final spellings in *y* or *ee*, but apparently a desire to retain the foreign appearance favored the final *i* spellings.

The alternation of *i* and *y* occurs with suffixation in the following cases (for exceptions, see *Webster's New Collegiate Dictionary*, 1956, pp. 1145-47):

1. When *ing* is added to words ending in *ie*, *e* is dropped and *i* is changed to *y*, e.g., *die*:*dying*, *tie*:*tying*.

2. When a suffix not beginning with *i* is added to a base ending in *y*, *y* changes to *i*, e.g., *icy*:*iciest*, *mercy*:*merciless*.
3. A few words which end in *vowel* + *y* irregularly change *y* to *i* before some suffixes, e.g., *day*:*daily*, *lay*:*laid*.

i and *y* also alternate as parts of the compound vowel units *ai* (*ay*), *ei* (*ey*), and *oi* (*oy*), the *y* alternate appearing before vowels and word juncture and the *i* alternate appearing in all other cases. Exceptions to this pattern are the following:

geyser	An Icelandic borrowing, taken from the name of a hot spring in Iceland.
oyster	A ME adoption of OF *oistre*, *uistre*. Judging from the OED entries for this word, *oy* was the only spelling commonly employed. *oi* is rare, a phenomenon for which I can offer no satisfactory explanation.
paranoia	A late nineteenth century Latin adoption. The *oia* spelling was probably retained because of its similarity to other medical terms ending in *ia*, e.g., *meglomania*, *malaria*, *anemia*, *dementia* and *hysteria*.

sequoia	A Latin-like term modeled after *Sequoiah*, a Cherokee Indian name.
moiety	A ME adoption of OF *moite*, *moitie*. The earlier form *moitie* was replaced by *moiety* in the last half of the seventeenth century. The middle vowel in the pronunciation is probably excrescent.

u and *w*

As consonant spellings, *u* and *w* are, for the most part, in complementary distribution. *u* appears as a spelling for :w: in the environments *gu*, *su*, *qu* (initial and medial), *ju*, *nu* (all medial), *pu* and *cu* (initial). *w* occurs in all other cases, except for the three forms *suave*, *suede*, and *suite*. Following is an exhaustive listing of occurrences of these spellings in the corpus used for this study.

gu

anguish
distinguish
extinguish
jaguar
language
languid
lingual
linguist
penguin
sanguine
unguent

su

assuage
persuade

pu

pueblo

cu

cuirass
cuisine

ju

marijuana

nu

ennui

u and w also alternate as the second parts of the compound vowel spellings which begin with a, e, and o. The w variant generally appears before another vowel spelling and in morpheme final positions, while the u variant occurs in all other positions. This pattern, however, is far from regular.

Examples of Regular Spellings

au	aw	eu	ew
auction	awe	eucharist	ewe
audit	draw	eulogy	ewer
augur	gnaw	feud	brewer
applaud	saw	neuter	shrewish
fault	straws	pneumatic	flew
trauma	thaw	rheumatism	threw

ou	ow
ounce	coward
oust	however
out	power
compound	allow
south	cow
trousers	vow

Exceptions

a. aw in initial and medial position before a consonant

awkward
awl
awning
dawn
scrawl
spawn

b. ew in medial position before a consonant

newt
pewter
lewd
shrewd

c. ou in final position

bayou
caribou
you
bijou
thou

d. ow in initial and medial position before a consonant

owl
own
crowd
drowse
fowl
town

u in final position in English orthography is rare. Besides the examples given above, the only other common examples are the following.

beau
bureau
chateau
plateau
portmanteau
trousseau
adieu
lieu
flu
gnu
impromptu
menu
parvenu
virtu

ous and ou

The suffix ous becomes os before ity. Thus, viscous: viscosity, curious:curiosity, generous:generosity, porous: porosity, monstrous:monstrosity.

i and *e*

In the suffix *ity*, *i* becomes *e* when preceded by *i*. Thus *society*:*propriety*, *piety*:*satiety*, *contrariety*:*anxiety*, *variety*:*sobriety*.

RESULTS - 2

Consonant Correspondences and Alternations

A Summary

b

Correspondences

1. b in debt, doubt and subtle corresponds to :0:.[1]
2. b elsewhere corresponds to :b:.

Regular Alternations

1. Initial :b: in :bd-: and morpheme final :b: in :-mb: correspond to :0:. (Bombard is classed as a non-productive alternation with bomb; see below.)
2. In the medial combination bilabial stop + bilabial stop, the first stop is lost. Cf., for example, cupboard, subpoena, clapboard, raspberry.

Non-Productive Alternations

:b: ~ :0:

bombard	bomb
crumble	crumb
debit	debt
indubitable	doubt
number	numeric
thimble	thumb

c

Correspondences

1. c in cello and concerto corresponds to :č:.

---

[1] Morphophonemic entities are enclosed in colons.

2. c in victual, czar, and indict corresponds to :0:.
3. c before the spellings i, y, e (and in facade) corresponds to :s:.[1]
4. c in all other positions corresponds to :k:.

Regular Alternations

1. :sj: before an unstressed vowel and in sure and sugar corresponds to :š:.

Non-Productive Alternations

:k: ~ :0:

corpuscular	corpuscle
muscular	muscle
suspect	suspicion

:k: ~ :s:

lyric	lyricism
music	musician[2]

ch

Correspondences

1. ch in fuchsia, yacht, drachm corresponds to :0:

---

[1] Except in sceptic, an alternate spelling of skeptic, where c corresponds to :k:.

[2] :s: is palatalized by :j: to :š:.

2. ch corresponds to :k: before l, n, r, and before final ar (in polysyllabic words) and in the following words:

ache
catechism
chameleon
chaos
character
chasm
chemi-
chiropodist
choir
choler
cholera
choral
chord
chorus
dichotomy
echo
epoch
eucharist
eunuch
hypochondria
lichen
machination
mechanic
orchestra
orchid
pentateuch
psych-
stomach

3. ch corresponds to :š: in the following words:

cache
chagrin
chaise
chalet
challis
chamois
champagne
champaign
chandelier
chanty
chaperon
charade
charivari
charlatan
chassis
chauffeur
chauvinism
chef
chemise
cheroot
chevalier
cheviot
chevron
chicanery
chiffon
chivalry
chute
cliche
crochet
echelon
machine
mustache
ricochet
sachet
stanchion

4. In all other cases, ch corresponds to :č:.

Regular Alternations

None. (On :k:~:č:, see Non-Productive Alternations under k.)

Non-Productive Alternations

None.

d

Correspondences

d regularly corresponds to :d:. (The identity of the past tense and participle marker :d: must be retained at the morphophonemic level so that the alternation of :d: and :t: can be handled properly.)

Regular Alternations

1. The past tense and participle marker :d: becomes :t: after all unvoiced final consonants except :t:.
2. :d: corresponds to :0: between :n: and another consonant.[1]
3. :dj: palatalizes to :ǰ: before unstressed :u:.

---

[1] This rule applies generally to hands, pounds, stands, handful, grandmother, but less frequently to handstroke and landslide.

Non-Productive Alternations

:d: ~ :s:

defend	defense
expend	expense
respond	response

:d: ~ :0:

sound	sonant
spindle	spin

f

Correspondences

1. f in of corresponds to :v:.
2. In all other cases, f corresponds to :f:.

Alternations

See Results - 4.

g

Correspondences

1. g corresponds to :0: in seraglio.
2. g corresponds to :ž: in the following words (some have alternate :ǰ: pronunciations):

bourgeois
camouflage
corsage
garage

lingerie
loge
massage
mirage

prestige
protege
regime
rouge
sabotage

3. g in gaol corresponds to :ǰ:.

4. g corresponds to :g: in the following:

anger
auger
beget
begin
begirth
bogey
conger
eager
finger
fogey
forget
gear
get

geyser
gibbon
giddy
gift
gig
giggle
gild
gill
gilt
gingham
girder
girdle
girl

girth
give
gizzard
hunger
linger
meager
monger
porgy
renege
target
tiger
together
yogi

5. g before e, i, and y corresponds to :ǰ:.

6. g otherwise corresponds to :g:.

Regular Alternations

1. Morpheme final :-ŋg: becomes :ŋ: in all cases except before the comparative and superlative morphemes, where it remains as :-ŋg:. Thus, :strɔŋ: : :strɔŋger:, but :sɪŋ: : :sɪŋər:.

2. Word initial :gn-: becomes :n: (cf. gnostic and agnostic).

3. Morpheme final :-gn: and :-gm: become :n: and :m:. Thus sign :saɪn: : signal :sɪgnəl:.

4. :g: in all other cases corresponds to :g:.

Non-Productive Alternations

:g: ~ :ǰ:

legal	legislate

:ǰ: ~ :k:

laryngeal	larynx

gh

Correspondences

1. *gh* in *hough* corresponds to :k:.
2. *gh* corresponds to :f: in

clough
cough
draught
enough
laugh
rough
slough
sough
tough
trough

3. Initial *gh* corresponds to :g:, e.g., *ghastly*, *ghetto*, *ghost*, *ghoul*.
4. *gh* corresponds to :g: in

aghast
burgh
dinghy
sorghum
spaghetti

5. In all other cases, *gh* corresponds to :ø:.

Alternations

None.

h

Correspondences

1. Initial *h* in *heir*, *herb*, *honest*, *honor*, *hour* and *humor* corresponds to :∅:.
2. Post vocalic *h* before a consonant or word juncture is a marker for the preceding vowel and therefore corresponds to :∅:, e.g., *fahrenheit*, *hurrah*.

Regular Alternations

1. :h: preceded by a consonant corresponds to :∅: (in most English dialects), cf., e.g., *philharmonic*: *harmonic*.
2. Intervocalic :h: when preceded by a stressed vowel, corresponds to :∅:, cf., e.g., *vehicle*:*vehicular*.
3. In all other cases :h: corresponds to :h:.

Non-Productive Alternations

None.

j

Correspondences

j corresponds to :ǰ:, except in *bijou* (*j*→:ž:), hallelujah, (*j*→:j:), and *marijuana* (j→:∅:).

Alternations

None.

k

Correspondences

k regularly corresponds to :k:.

Regular Alternations

1. Word initial :kn-: becomes :n:, as in *knee* and *knowledge*. (In *acknowledge* :k: is retained. This is the only example of the :kn: ∼ :n: alternation. The parallel alternation of :gn: ∼ :n: occurs in *agnostic*:*gnostic*.)
2. :k: in all other cases corresponds to :k:.

Non-Productive Alternations

:k: ∼ :č:

leak	leach
speak	speech
wreak	wretch

l

Correspondences

1. The first *l* in *colonel* corresponds to :r:.
2. In *could*, *should*, and *would* *l* corresponds to :∅:.
3. In all other cases *l* corresponds to :l:.

Regular Alternations

1. In morpheme-final :æ/a: + :l: + :k/m/v/f:, :l: corresponds to :∅:, e.g., *half*, *calm*, *chalk*, *halves*.

2. In prejunctural position, consonant + *l* corresponds to consonant + :əl: (syllabic l). cf. *simply*:*simple*.
3. In all other cases :l: corresponds to :l:.

Non-Productive Alternations.

None.

m

Correspondences

1. In *accompt* and *comptroller*, *m* corresponds to :n:.
2. Otherwise, *m* corresponds to :m:.

Regular Alternations

1. Initial :mn-: corresponds to :n:.
2. After :ð: and :z:, prejunctural :m: becomes :-əm: (syllabic :m:), e.g., *rhythm*, *spasm*, *schism*.

Non-Productive Alternations

1. *com*-, *con*-, *co*-, e.g., *complex*, *conclude*, *coeval*.
2. *sym*-, *syn*-, *syl*-, e.g., *symphonic*, *synonym*, *syllogism*.

n

Correspondences

1. :n: in *kiln* corresponds to :∅:.
2. Otherwise *n* corresponds to :n:.

Regular Alternations

1. :n: before :k: and :g: becomes :ŋ: except in the un-

stressed prefix - morphemes *con-*, *syn-*, *in-*, where it remains as :n: (with considerable idolectal and dialectal variation).

2. :n: in all other cases corresponds to :n:.

Non-Productive Alternations

:n: ~ :∅:

passenger	passage
messenger	message

p

Correspondences

1. *p* in *corps*, *coup*, and *receipt* corresponds to :∅:.
2. *p* in all other cases corresponds to :p:.

Regular Alternations

1. Initial :p: in :pn-, ps-, pt-: corresponds to :∅:.
2. In the medial sequence bilabial stop + bilabial stop, the first stop is lost.

Non-Productive Alternations

:p: ~ :∅:

reception	receipt[1]

---

[1] Kurath (1964, p.39) lists *(un)kempt* ~ *comb* as an example of the :p: ~ :∅: alternation.

ph

Correspondences

ph corresponds regularly to :f:, except in phthisic where it corresponds to :ø:.

Alternations

None.

q

Correspondences

q corresponds regularly to :k:.

Alternations

None.

r

Correspondences

r in all cases corresponds to :r:.

Alternations

None.[1]

rh

Correspondences

rh corresponds regularly to :r:.

---

[1] The dialectal variations of r are beyond the scope of this report.

Alternations

None.

s

Correspondences

The rules which follow are not exhaustive, and admit numerous exceptions. Considerable work remains to be done with s.

1. s corresponds to :ø: in

   aisle
   corps
   demesne
   island
   isle
   apropos
   bourgeois
   challis
   chamois
   chassis
   debris
   rendezvous
   velours

2. Final inflectional s corresponds to :s:. (The identity of inflectional s must be retained on the morphophonemic level.)
3. Final s after a voiced consonant spelling and in as, has, his, is, was, corresponds to :z:.
4. Final s in all other cases corresponds to :s:.
5. Initial s corresponds to :s:.
6. Medial s before or after a voiceless consonant spelling and in the combinations -ase, -ease, -vowel + rse corresponds to :s:.
7. s in all other cases corresponds to :z:.

Regular Alternations

1. After :s, z, š, ž, č, ǰ:, morphemic :s: becomes :-ɪz:; after any other voiced morphophoneme it becomes :z:; otherwise, it remains as :s:.
2. On other :s: ~ :z: alternations, see Results - 4.

Non-Productive Alternations

See Results - 4.

sh

Correspondences

sh corresponds regularly to :š:.

Alternations

None.

t

Correspondences

1. *t* corresponds to :∅: in *depot*, *debut*, *hautboy*, *savant*, *mortgage*, and in words ending in ⟨-et⟩ with stress on the last syllable (e.g., *buffet*, *valet*).
2. *t* in *equation* corresponds to :z:.
3. In the combination *t* + *vowel* + *vowel*, *t* corresponds to :s: when not preceded by *s* or *x* (cf. *nation*:*bastion*.)
4. *t*, otherwise, corresponds to :t:.

Regular Alternations

1. On the palatalization of :sj: to :š:, :zj: to :ž: and :tj: to :č:, see Weir and Venezky, 1965, pp. 44-48.
2. Morpheme final :-st: becomes :s: where followed by :-en: : *-en* (cf. *haste*:*hasten*.)
3. In the medial cluster, :-stl-: and :-ft:, :t: corresponds to :∅:.
4. :t: otherwise corresponds to :t:.

Non-Productive Alternations

:t: ∼ :č:

cent	century
cult	culture
digest	digestion

:t: ∼ :š:

act	action
permit	permission
part	partial

:t: ∼ :ž:

convert	conversion

:t: ∼ :s:

-ant	-ance (e.g., elegant:elegance)

th

Correspondences

1. *th* in *thyme* corresponds to :t:.
2. *th* in *isthmus* and *asthma* corresponds to :∅:.
3. Initial *th* in functors corresponds to :ð:
4. Medial *th* in the morpheme-final clusters -*the* and -*ther* (except for *ether*) corresponds to :ð:.
5. In *ether* and all other cases, *th* corresponds to :θ:.

Alternations

See Results - 4.

u

Correspondences

1. Forms ending in -*que*, -*quest*, -*quette*, and -*quee* have *u* corresponding to :∅:. In *quay*, *queue*, *liquor*, *piquant*, *turquoise*, *u* corresponds to :∅:.
2. Otherwise, consonantal *u* corresponds to :w:.

Alternations

None.

v

Correspondences

v corresponds regularly to :v:.

Regular Alternations

See Results - 4.

Non-Productive Alternations

1. :v: ~ :w:

   vine, vinter, wine

2. :v: ~ :f:

   see Results - 4.

w

Correspondences

1. w in answer, sword, two and toward corresponds to :0:.
2. w in initial position and after a consonant corresponds to :w:.
3. Post-vocalic w in award, aware, bewail, and beware corresponds to :w:.
4. In all other cases post-vocalic w is part of a vowel spelling.

Regular Alternations

None.

Non-Productive Alternations

see v.

wh

Correspondences

1. wh corresponds to :h: in who, whole, whom, whoop, whore, whose.
2. Otherwise, wh corresponds to :hw:.

Regular Alternations

None.

Non-Productive Alternations

None.

x

Correspondences

1. Initial x corresponds to :z:.
2. Final x corresponds to :ks:.
3. Medial x corresponds to :k,s,:.

Regular Alternations

1. :k,s,: becomes :gz: in the following environments:
   a. Intervocalic with stress not on the preceding syllable.[1]
   b. Between a vowel and :h: (e.g., exhibit).
2. Otherwise, :k,s,: becomes :ks:.

---

[1] Exceptions are doxology, exit and exile when pronounced with /gz/, and exhale.

y

Correspondences

1. Initial y corresponds to :j:.
2. Medial y in beyond and lawyer corresponds to :j:.
3. Post-vocalic y in all other cases is part of a vowel spelling.

Regular Alternations

None.

Non-Productive Alternations

None.

z

Correspondences

1. z in rendezvous corresponds to :∅:.
2. zz in mezzo and pizza corresponds to :ts:.
3. Otherwise, z corresponds to :z:.

Regular Alternations

1. Final :z: after a voiceless consonant corresponds to :s:.
2. Otherwise, :z: corresponds to :z:.

Non-Productive Alternations

None.

RESULTS - 3

Vowel Correspondences and Alternations

A Summary

Primary Vowel Patterns

Introduction

The vowel spellings *a*, *e*, *i*/*y*, *o*, *u*, called primary vowel spellings in this paper, carry the major burden of vowel representation in the current orthography.[1] They occur in all positions and have a vast complexity of morphophonemic correspondences and alternations which reflect an even more complex history. When viewed from the direct spelling-to-sound standpoint, the patterns for these units reveal no regularity. *o* corresponds to seventeen different sounds, *a* to ten, *e* to nine, and the combined group to forty-eight. When the morphemic structure and consonant environments of the words in which these units appear are analyzed, however, a single major pattern emerges from which regular sub-patterns can be derived. Exceptions still remain, but the underlying pattern is so dominating that the exceptions, which were once the rule, become mere oddities, begging for historical justification. In the discussion which follows the major pattern for the stressed vowels is introduced in a general form and then refined through the introduction of its regular subpatterns, alternations and exceptions.[1]

---

[1] Henceforth *i* will stand for both *i* and *y*.

Major Pattern

Each of the primary vowel units corresponds regularly to two different morphophonemes, a checked one and a free one, according to the morphemic structure of the word in which it occurs and the consonant and vowel units which follow it. These correspondences are shown in Table 1 below.

---

1. Word stress patterns play a significant role in the relationship of spelling to sound, especially in the correspondences of the primary vowel spellings. With a few exceptions, only stressed vowels are treated in this paper. Three levels of stress, all introduced on the morphophonemic level, are assumed: primary, secondary, and tertiary (unstressed). While some linguists have claimed without qualification that English word stress is "predictable," no extensive analysis of this topic has ever been published. Two recent publications, however, indicate that work is being done in this area. See George S. Waldo, "The significance of accentuation in English words," Proceedings of the ninth international congress of linguists (The Hague, 1964), pp. 204-210, and Roger Kingdon, The groundwork of English stress, (London, 1958).

TABLE 1

MAJOR PATTERN FOR PRIMARY VOWELS

Spelling	Free alternate	Checked alternate
a	:e:	:æ:
	sane	sanity
	mate	mat
	ration	rattle
e	:i:	:ɛ:
	athlete	athletic
	mete	met
	penal	pennant
i	:ai:	:ɪ:
	rise	risen
	malign	malignant
	site	sit
o	:o:	:a:
	cone	conic
	robe	rob
	posy	possible
u	:ju:[1]	:ə:
	induce	induction
	rude	rudder
	lucre	luxury

[1] The retention or elision of :j: before :u: is handled as a morphophonemic process.

In monomorphemic words a primary spelling unit corresponds to its free alternate when it is followed by (1) a functionally simple consonant unit which in turn is followed by another vowel unit (including final *e*) or (2) a functionally simple consonant unit, followed by *l* or *r*, and then another vowel unit (including final *e*).[1] It corresponds to its checked alternate in the remaining cases, i.e., when followed by (1) a functionally compound consonant unit, e.g., *x*, *dg*, (2) a cluster of consonant units, e.g., -*nn*, *lth* or, (3) a word-final consonant unit or units. Examples of these correspondences are shown in Table 2 below. The column numbers correspond to the numbered qualifications in the sentences above.

---

[1] The difference between monomorphemic and polymorphemic words, a difference unfortunately neglected in the teaching of reading, is too complex to be discussed adequately here. For the prediction of sound from spelling in a large number of words, however, the distinction is crucial. What are involved, primarily, are the morphophonemic alternations which occur with suffixation, as in /ərbén/ /ərbǽnɪtɪ/, /kón/:/kánɪk/. The best linguistic material on English suffixation is found in Stanley S. Newman, "English suffixation: a descriptive approach," *Word* 4.24-36 (1948). For a pedagogical approach to this subject, see Edward L. Thorndike, *The teaching of English suffixes* (New York, 1941).

TABLE 2

EXAMPLES OF PRIMARY VOWEL CORRESPONDENCES FOR SELECTED ENVIRONMENTS

Spelling	Free alternate		Checked alternate		
	1	2	1	2	3
a	canine	ladle	badge	saddle	sat
e	median	zebra	exit	antenna	ebb
i	pilot	microbe	chicken	epistle	hitch
o	vogue	noble	pocket	cognate	sod
u	dubious	lucre	luxury	supper	rug

The correspondences for these vowel spellings in polymorphemic words depends not only upon the graphemic environment, but also in many cases upon the morphemic structure of the word. Correspondences for some polymorphemic words have been discussed previously (see Weir and Venezky, 1965 , pp. 42-44.)

Sub-patterns

The two most important sub-patterns which can be derived from the major pattern are the final e pattern and the geminate consonant pattern. Examples of these are shown in Table 3 below; complete analyses are given in the two following sections.

TABLE 3

EXAMPLES OF FINAL *e* AND GEMINATE CONSONANT PATTERNS

Spelling	Final *e* pattern	Geminate consonant pattern
a	r*a*te -- r*a*t	*a*nal -- *a*nnals
e	m*e*te -- m*e*t	P*e*ter -- p*e*tter
i	s*i*te -- s*i*t	d*i*ner -- d*i*nner
o	p*o*pe -- p*o*p	c*o*ma -- c*o*mma
u	c*u*te -- c*u*t	s*u*per -- s*u*pper

Final *e* Pattern

While the final *e* pattern applies primarily to monosyllabic words, it also holds for many polysyllabic words, even when the vowel before the final *e* is unstressed, as in *microbe*, *decade*, *schedule*, *volume*, *placate*. Besides the patterns mentioned above (*vowel* + *consonant* + *e*, *vowel* + *consonant* + *le*, *vowel* + *consonant* + *re*), the environment *vowel* + *ste* is also part of the final *e* pattern, as in *baste*, *chaste*. Examples of regular correspondences are shown below, followed by an exhaustive list of exceptions for stressed vowels.

TABLE 4

EXAMPLES OF THE FINAL e PATTERN

a→:e:	i→:ai:	u→:ju:	e→:i:	o→:o:
bake	cycle	cube	accede	cove
decade	domicile	duke	impede	erode
fable	five	mule	obsolete	globe
gage	prize	produce	scheme	joke
haste	profile	resume	serene	mediocre
shake	sublime	secure	theme	smoke

Irregular Correspondences for the Final e Pattern

1. a corresponds to :æ: in bade, forbade, have, morale.

   a corresponds to :a: in are, barrage, camouflage, corsage, facade, garage, massage, mirage, sabotage.

   Note. a in the ending -ate corresponds to :e: in verbs, but alternates to :ɪ: in nouns and adjectives. Cf. duplicate (vb.):duplicate (adj.,noun). Stress is also important in this alternation.

2. e corresponds to :ɛ: in allege, ere, there, treble, where.[1]

   e corresponds to :ɪ: in clientele, renege.

---

[1] In the following words final e is not a marker, but a relational unit, (1) corresponding to :ɪ: : adobe, coyote, epitome, extempore, facsimile, finale, nike, hyperbole, sesame, simile, ukulele, recipe, (2) corresponding to :e: : cafe, protege.

e corresponds to :ə: in were,

e corresponds to :e: in fete.

3. i corresponds to :i: in the following words:

bastile
caprice
castile
elite
machine
marine
police
prestige
ravine
regime
routine
sardine
tangerine
valise

i corresponds to :ɪ: in give and live.

Note. The spelling i is highly irregular in the ending -ine when unstressed. Cf. canine, asinine: examine, famine.

4. o corresponds to :ə: in above, come, done, dove, glove, love, none, shove, some.

o corresponds to :u: in lose, move, prove, whose.

o corresponds to :ɔ: in gone.

Vowel Spellings before Geminate Consonants

The primary spellings a, e, i, o, u occur frequently before geminate consonants, the digraph (secondary) spellings, rarely.[1] (*Braille*, *chauffeur*, and *trousseau* are the only examples from the corpus used for this study.) Before geminate consonant clusters, primary vowel spellings correspond to their checked alternates, with the following exceptions:

---

[1] Except for the Spanish borrowing *llama*, geminate consonants occur only in medial and final position in English spelling. The most common final clusters are *ff*, *ll*, and *ss*, as in *fluff*, *mill*, and *toss*. Rare final clusters are *bb*, *dd*, *gg*, *nn*, *rr*, *tt*, and *zz*, which occur only in the words *ebb*; *add*, *odd*; *egg*; *inn*; *err*, *burr*, *purr*, *whirr*; *boycott*, *butt*, *putt*, *watt*; *buzz*, *fizz*, *fuzz*, *jazz*, *razz*. In medial position the following are common:

bb	cabbage, lobby	*mm*	comment, summit
cc	broccoli, occupy	*nn*	annals, dinner
dd	addict, eddy	*pp*	copper, happy
ff	coffee, giraffe	*rr*	error, surrey
gg	luggage, nugget	*ss*	blossom, gossip
ll	fillet, pillow	*tt*	button, clutter

vv and *zz* also occur medially, but only in *flivver*, *navvy*, *blizzard*, *dizzy*. No other geminate clusters occur, although *dg*, *ck*, and *tch* can be considered as geminate substitutes.

Consonant gemination also occurs when certain suffixes are added to words ending in a single consonant. For example, *run*: *running*, *forbid*:*forbidding*, *abet*:*abetted*. The orthographic rule is usually stated as follows:

> Monosyllables and words accented on the last syllable, when ending in a single consonant, *b*, *d*, *f*, *g*, *k*, *l*, *m*, *n*, *p*, *r*, *s*, *t*, *v* preceded by a single vowel, double the

---

consonant before adding a termination beginning with a vowel or the suffix -*y*. *Webster's*, 1956, p. 1145. Exceptions to this rule are plentiful and usage is still unsettled on whether or not to double final *l*. For a colorful survey of eighteenth and nineteenth century grammatical literature on this topic, see G. Brown *Grammar of English Grammar* , (4th ed. New York, 1859) pp. 198-200.

The non-gemination of *v* and *th* has led to a large number of exceptions to the major pattern for the correspondences of the primary vowel spellings. *Cover*, *bevel*, *level*, *river*, *brother*, *mother*, *other*, for example, have vowel spellings corresponding to checked alternates in environments which indicate free alternates. To indicate the checked alternate, *v* and *th* would have to be geminated, but the graphotactical patterns of English exclude the doubling of these units.

In *prison* a slightly different problem exists. While *ss*, which is needed to make the correspondence *i*-:ɪ:, is allowed, it generally corresponds to :s: in medial position, as in *blossom*, *gossip*, and *lasso*, so it could not be employed where *s* corresponds to :z:. The present use of *ss* is derived from Old French orthography where "intervocalic *ss* served to distinguish voiceless s from voiced s (=z)." Alfred Ewert, *The French language* (London, 1933), p. 113.

1. *a* in *mamma* corresponds to :a: and *a* in *marshmallow* corresponds to :ɛ:.
2. *o* in *across*, *albatross*, and *boss* corresponds to :ɔ:.
3. *o* in *boggy*, *gross*, and before final *ll* corresponds to :o:.
4. *u* in *butte* corresponds to :ju: and *u* in *pudding* and *pussy* corresponds to :u:.

TABLE 5

EXAMPLES OF REGULAR CORRESPONDENCES BEFORE GEMINATE CONSONANTS

a→ :æ:	*e*→ :ɛ:	*i*→ :ɪ:	*o*→ :a:	*u*→ :ə:
abbess	appellate	artillery	accommodate	button
accent	beggar	auxillary	collar	funnel
apple	bellow	blizzard	college	funny
babbittry	cellar	bacillus	comma	hullabaloo
cabbage	cheddar	cribbage	commerce	hummock
callow	dilemma	issue	dollar	mummy
fallacy	ebb	shrill	hobby	puddle
flabby	fellow	symmetry	hollow	pummel
grammar	kennel	vanilla	lobby	rubber
happen	lesson	village	sonnet	shutter
mammal	message	willow	toboggan	supper
rattle	tennis	wriggle	toggle	tunnel

Alternations Based Upon Primary Vowel Spellings

The major pattern for the primary vowel spellings in stressed positions depends upon two basic features: environment, which was discussed in the previous sections, and morphemic structure, which is discussed briefly in this and the following sections. Morphemic structure forms the basis for describing the morphophonemic alternations based upon the primary vowel spellings. For example, the word *sanity*, if considered solely on the basis of the rules given in the preceding section, would be an exception to the major pattern since *a* before a simple consonant unit followed by a vowel corresponds to its checked rather than its free alternate. If viewed, however, in relation to the sequences, *sane*:*sanity*, *humane*:*humanity*, and *urbane*: *urbanity*, another regular feature can be seen. If we start with the forms *sane*, *humane*, and *urbane*, then regular rules can be written for changing the free alternate :e: to the checked alternate :æ: when the suffix -*ity*: :-ɪtɪ: is added.

This rule also holds for the spellings *e*, *i* and *o*, as can be seen from the following examples:

e	extr*e*me	extr*e*mity
	obsc*e*ne	obsc*e*nity
	ser*e*ne	ser*e*nity
	supr*e*me	supr*e*mity

i	asin*i*ne	asin*i*nity
	dev*i*ne	div*i*nity
	mal*i*gn	mal*i*gnity
	sen*i*le	sen*i*lity
o	friv*o*lous	friv*o*lity
	medi*o*cre	medi*o*crity
	prec*o*cious	prec*o*city
	verb*o*se	verb*o*sity

(Morphophonemic alternations based upon stressed *u* are rare in Modern English, the most common being those which occur in *assume*:*assumption*, *conduce*:*conduction*, *presume*:*presumption*, *reduce*:*reduction*. Even with loss of stress *u*: :ju: tends not to change to :ə:. Thus, *compute*:*computation*, *execute*:*execution*, *usurer*:*usurious*, *utilize*:*utility*.)

While a complete survey of vowel morphophonemics is beyond the scope of this paper, some of the more common alternations are presented below.

1.	-*ic*:	free --	checked
	a	*a*ngel	*a*ngelic
		st*a*te	st*a*tic
	e	athl*e*te	athl*e*tic
		hygi*e*ne	hygi*e*nic
		m*e*ter	m*e*tric

	i	cycle	cyclic
		mime	mimic
		paralyze	paralytic
	o	cone	conic
		neurosis	neurotic
		phone	phonic
2.	-ion	free --	checked
	a	compare	comparison
	e	concede	concession
		convene	convention
		descrete	descretion
	i	collide	collision
		decide	decision
		provide	provision

Secondary Vowel Patterns

Introduction

The secondary vowel spellings differ from the primary vowel spellings in several important ways. First, they occur less frequently and have a more limited distribution. None appears commonly before geminate consonant clusters; some like ai, au, ei, and eu rarely occur in word-final position; others, like ie and oa rarely occur in word initial position. Second, while each primary vowel spelling has two basic correspondences, according to the graphemic environment and the morphemic composition of the word in which it occurs, each secondary vowel spelling generally has a single major correspondent. Third, the morphophonemic correspondences based upon the secondary vowel spellings tend not to alternate in quality with reduction in stress. Compare, for example, the first vowels in neutral:neutrality, cause:causation with those in melody:melodious, potent:impotent.

Historically, the secondary spellings also differ from the primary ones. Primary vowel spellings are found in the earliest English records; their correspondences can be traced through a complicated chain of sound changes from Old English to the present time. Most secondary vowel spellings, on the other hand, were introduced during the late Middle English period and consequently have been involved in considerably fewer sound changes. For a complete summary of the secondary vowel spellings, see Weir and Venezky, 1965, pp. 28-38.

RESULTS - 4

The Morphophonemics of Voiced-Voiceless Alternations in American English Fricatives

The morphophonemic alternations :s~z:, :f~v:, :θ~ð:, :š~ž: reveal a number of problems in the construction of a model for spelling-to-sound correspondences. These problems, though encountered here in traversing from the orthography toward sound, also occur in the construction of other models for English morphophonemics, so that the considerations presented below have a much wider significance than in spelling-to-sound studies alone. In the discussion which follows we assume that spellings have been mapped onto an initial morphophonemic level and that our next objective is to describe morphophonemic alternations in terms of regular and irregular rules. In this section we will assume that an alternation that is both productive and high frequency is labeled regular. The specific problems which are discussed in relation to the voiceless-voiced alternations are

(1) what generalizations can be made about these alternations and

(2) what relevant differences exist among the non-productive (irregular) patterns.

Whether or not an alternation is marked in the orthography is relevant for both reading and spelling. This fact will be mentioned for each pattern, but not discussed thoroughly.

Summary of the Alternations

I Noun plural :s~z (~ɪz): cats:dogs(:horses) (unmarked)

II Noun-adjective

a. :š~ž: *luxury*:*luxurious* (unmarked)

b. :s~z: *louse*:*lousy* (unmarked)

c. :θ~ð: *north*:*northern* (unmarked)

III Verb-verb

:s~z: *sound*:*resound* (unmarked)

IV Noun-verb (infinitive)

a. :f~v: *belief*:*believe* (marked)

b. :θ~ð: *breath*:*breathe* (marked)

c. :s~v:

1. *house* (n.):*house* (v.) (unmarked)
2. *glass*:*glaze* (marked)

V Noun singular-noun plural

a. :f~v: *wife*:*wives* (marked)

b. :θ~ð: *bath*:*baths* (unmarked)

c. :s~z: *house*:*houses* (unmarked)

VI Verb (past)-verb (present)

:f~v: *left*:*leave* (marked)

VII Noun-noun

:θ~ð: *smith*:*smithy* (dialectal) (unmarked)

VIII Ordinal-cardinal

:f~v: *fifth*:*five* (marked)

Of these alternations, only a few are invariant. The noun plural (I) certainly is. It is a phonetically conditioned alternation which is describable by a single rule. It is generally not marked in the spelling, as shown by the following plurals, all of which add only *s* to the singular:

houses	[háuzɪz]
boys	[bɔ́ɪz]
cats	[kǽts]

(Where -*es* is added for the plural, the alternate :ɪz: is marked.)

Class VIII, the ordinal-cardinal alternation, is also invariant since the only two members which end in voiced :v: in the cardinal have unvoiced :f: in the ordinal, viz., *five*:*fifth*, *twelve*:*twelfth*. One rule for the generation of the ordinal forms from the cardinal forms, therefore, can be stated as

final :v: → :f:.

It is also possible to state this rule as follows:

voiced fricative → unvoiced fricative

because no other cardinals end in a voiced fricative. Such a generalization, however, is probably not acceptable unless there are other major patterns in the orthography which show a voiced-voiceless alternation in the fricatives. The existence of such patterns beyond the noun plural are explored in the remainder of this section.

Class IV b, the :θ~ð: alternation in noun-verb, is nearly invariant, but since it has so few examples, there is a temptation to base its classification upon those of the other class IV alternations. The nouns which end in :θ: and which can also be used as verbs are the following:

:θ~ð:		:θ~θ:	
n.	v.	n.	v.
breath	breathe	froth	froth
cloth	clothe	earth	(un)earth
mouth	mouth		
swath	swathe		
teeth	teethe		

The :f~v: alternations for noun-verb are not nearly so regular, as the following lists demonstrate.

:f~v:		:f~f:	
n.	v.	n.	v.
belief	believe	brief	brief
grief	grieve	dwarf	dwarf
half	halve	leaf	leaf
knife	knive	sniff	sniff
life	live	spoof	spoof
proof	prove	strafe	strafe
relief	relieve	stuff	stuff
shelf	shelve	whiff	whiff
thief	thieve		

It would be stretching a point considerably to state that in general nouns ending in :f: changed :f: to :v: in forming the corresponding infinitive. At best, it can be said that there is a class of nouns (which must be enumerated in the lexicon attached to the model) in which the voiceless-voiced alternation occurs for this production.

The case for a :s~z: alternation in this class is even more tenuous. *Abuse*, *grease* (dialectally), *house*, and *use* have this alternation, as do in an extended sense *brass* (*braze*), *glass* (*glaze*), and *grass* (*graze*). But there is an even longer list of forms like *class*, *curse*, *dress*, *eclipse*, *glimpse*, *mass*, *nurse*, *pass*, *press*, *stress*, and *witness* in which no such alternation occurs. In general, nouns ending in :s: do not change :s: to :z: in forming the corresponding infinitive.

On the basis of this evidence there is no justification for establishing a regular voiceless-voiced alternation for this class. This indicates, furthermore, that no special status should be assigned to the :θ~ð: alternation, even though it occurs in five of the seven possible cases. The only conclusion that can be drawn is that nouns ending in voiceless fricatives generally retain the voicelessness in forming the infinitive. There is, however, a productive subpattern in which the voiceless fricative is voiced.

Class II a which has only one example in the corpus used for this study, and probably not very many more in all of the English language, might also be classed as regular. But this depends upon whether the alternation :ks:~:gz: is classed as regular or not. If spelling is used as the starting point, then this is possible (but not necessarily desirable). Intervocalic *x* could be mapped into :$k_x s_x$: and then a stress rule applied to convert :$k_x s_x$: to :gz: (:$k_x s_x$:→:gz: when the primary stress is not on the preceding vowel, cf. *execute*:*exist*). :$k_x s_x$: must be distinguished from :ks: on the morphophonemic level since :ks: derived from *cc* does not become :gz: under any circumstances (cf. *access*:*accessible*, *accept*:*acceptable*). Under the classifications of *x* derived :ks: and *cc* derived :ks:, the :š~ž: alternation in *luxury*:*luxurious* is regular; the actual alternation, however, is not :š~ž:, but :s~z:. :š: and :ž: result from the palatalizations :sj:→:š: and :zj:→:ž:. This means first that the alternation :š~ž: is, from the standpoint of a generative model, actually the alternation :s~z: and second that it is regular only if the identity of *x*- derived :ks: is maintained until after stress is applied. If this identity is not maintained, then there appears to be no reason to call :ks:→:gz: regular. In lieu of the lack of other criteria to apply, frequency favors the invariantly voiceless cluster: *accent*:*accentuate*, *accept*:*acceptable*, *access*:*accessible*, *accident*:*accidental*. Therefore, the alternation in *luxury*:*luxurious* is not classed as regular here.

Class III could, and probably should be designated as regular. At present we have found only the three pairs

sign	design	
solve	dissolve	(absolve)
sound	resound	

The rule at first glance appears to be reasonably simple. Unfortunately, the voicing of :s: has not been worked out yet so at best we can say that we think that this alternation is regular, but can not prove it.

Class V is the largest of the remaining classes, containing the following pairs:

:s~z:		:f~v:		:θ~ð:	
house	houses	calf	calves	bath	baths
		elf	elves	cloth	clothes
		half	halves	lath	laths
		knife	knives	moth	moths
		leaf	leaves	mouth	mouths
		life	lives	path	paths
		loaf	loaves	wreath	wreaths
		self	selves	(Most of the plurals above vary between :θ: and :ð:. The dialect distribution for this alternation has not been fully determined.)	
		sheaf	sheaves		
		shelf	shelves		
		staff	staves		
		thief	thieves		
		wife	wives		
		wolf	wolves		
		beef	beeves	rare	
		hoof	hooves		
		scarf	scarves		
		wharf	wharves		
		turf	turves	(archaic)	

The lack of a :š~ž: alternation and the rarity of the :s~z: alternation eliminate the possiblity of a general voicing rule for final fricatives in forming the noun plural. This leaves the possibility of labeling the :f~v: and :θ~ð:

alternations as regular. But the regular label implies either that there is a phonetic or grammatical factor that will allow the prediction of the alternation, or that this alternation occurs with a high frequency. That the :f~v: or :θ~ð: alternations are not phonetically or grammatically predictable can be seen from the following items:

gulf	gulfs	breath	breaths	month	months
reef	reefs	width	widths	earth	earths
dwarf	dwarfs				
proof	proofs				

The frequency question also is answered in the negative, although the :f~v: alternation does occur in almost one-third of the nouns which end in final :f:. The remainder retain :f: in forming the plural. This would imply that the leaf:leaves alternation should be classed as irregular, just as the petrified alternation represented by smith:smithy. Our native speaker's intuition, however, is uncomfortable with this solution.

The smith:smithy alternation is not only uncommon, but also would be rarely invoked as a model for forming new English words. Leaf:leaves, on the other hand, appears more productive. We would suspect, for exampled, that voiced fricative plurals of such words as dwarf, fife, and roof occur in American speech. Furthermore, the formation of the noun plural from the noun singular is one of the most common forms of word formation in English. Formation of one noun from another by the addition of *-y* is much less productive.

The issue here is that a model that concerns itself strictly with regular and irregular categories fails to account for some features of the language. While much work remains to be done, we have tentatively decided to subdivide the irregular class into *productive* and *petrified* subclasses. At best we can say that alternations like *leaf*:*leaves* fit into the *productive* *irregular* class while those like *smith*:*smithy* fit into the *petrified* *irregular* group.

Finally, it should be noted that all :f~v: alternations are marked in the orthography by change in the functional unit while in the :s~z: alternations only the one represented by *glass*:*glaze* is marked, (also by a change of functional unit), and in the :θ~ð: alternations, only the noun-verb one is marked, and this by addition of the final *e* marker which is the only possible device for marking the voiced pronunciation of *th*.

A reclassification of these alternations, based upon the newly adopted classes, is shown below.

Regular

1.	noun plural	:s~z (~ɪz):	*cats*:*dogs*
2.	verb-verb	:s~z:	*sound*:*resound*
3.	ordinal-cardinal	:f~v:	*fifth*:*five*

Irregular

Productive

1.	noun-adjective	:θ~ð:	north:northern
2.	noun-verb	:θ~ð:	breath:breathe
		:f~v:	belief:believe
		:s~z:	house(n):house(v)
3.	noun-singular noun plural		
		:f~v:	wife:wives
		:θ~ð:	bath:baths
		:s~z:	house:houses

Petrified

1.	noun-adjective	:s~z:	louse:lousy
			luxury:luxurious
2.	noun-verb	:s~z:	glass:glaze
3.	verb (past)-verb (present)		
		:f~v:	left:leave
4.	noun-noun	:θ~ð:	smith:smithy

Conclusions and Suggestions for Further Research

Learning to read one's native language differs radically from learning to read a foreign language. For the latter task, that of reading a foreign language, the learner most frequently does not have any knowledge of the language he is to read and, in the majority of the cases, he desires to translate directly from writing to meaning. To the former task, however, the learner brings a reasonably adequate set of language habits, as evidenced by his ability to speak. Reading in this situation requires primarily the translation from written symbols to sound, a procedure which forms the basis of the reading process and furthermore, is probably the only language skill unique to reading. (Comprehension, for example, while a necessary criterion for reading, is a function of both speech and writing.) The primary concern of this paper is the teaching of this translation process; references hereafter to the teaching of reading, unless qualified otherwise, refer to spelling-to-sound translation only.

A good reader in the sense of this paper is one who can not only pronounce all of the words which he has been taught to read, but a high percentage of the new words which he encounters. Certainly a person who could do the former task -- that of pronouncing words he had seen before -- but displayed no ability to pronounce new words, would be classed as a deficient reader. Most literates form some spelling-to-sound generalizations regardless of the methods which they encountered in their initial reading instruction. What these generalizations are, how they develop, and how they differ from one literate to another is at present unknown.

The rules presented in this and our two previous reports represent an ideal system for translating from spelling to sound, formed by assigning equal weights to each of the words used in the original corpus. It is inconceivable that any human could without special effort (like reading these reports) arrive at the same rules. Type-token relationships probably are highly influential in the formation of such patterns. The pronunciations of initial *ch*, for example, are irregular according to our results since no consistent cues can be used for selecting among [š], [k], and [č]. The average literate, nevertheless, would most probably pronounce initial *ch* in a strange word as [č] simply because most of the frequently occurring words with initial *ch* are pronounced with [č].

Which of the major patterns are learned by literates is not known -- in fact, extremely little is known about the extent of generalization in this area. Furthermore, there has been so little discussion of this process that no clearly defined criteria for generalization have been formulated. In the *ch* pattern cited above, for example, what would good generalization be? To pronounce all new initial *ch* words with [č]? To pronounce X percent with [č], Y percent with [š], and Z percent with [k], where X, Y, and Z correlate to the frequency of known initial *ch* words with the different pronunciations? Or what? If the teaching of reading is to be placed on a more substantial foundation than it now occupies, then the nature and goals of reading generalization must be explored.

Once the reading habits of literates are known and attainable goals for reading established, some of the sacred cows of reading instruction could be re-examined. The first of these to undergo inspection should be the notion of sequencing materials from the simple to the complex, the heart of the Bloomfield approach and of almost all of the new "linguistic" approaches to teaching reading. While such an approach may, with sufficient experimentation, prove to be more efficient than any other possible approach, there is little linguistic or psychological support for it at present.

For example, understanding one of the most important spelling patterns, that of the correspondences for the primary vowel spellings (see page 57), requires differentiation of both graphemic environments and responses. The letter *a*, as an example, has two primary pronunciations in stressed position, /æ/ and /e/. The checked alternative, /æ/, occurs when *a* is followed by a final consonant or by a series of consonants, as in *rat* and *annals*. In addition, it occurs when *a* is followed by a single consonant plus one of several possible suffixes, like -*ity* (e.g., *sanity*). The free pronunciation, /e/, occurs when *a* is in the other graphemic environments, like *rate*, *anal*, and *sane*. What must be acquired for the proper pronunciation of *a* is the ability to differentiate the environments and suffixes; final consonant vs. consonant plus final *e* (*rat*:*rate*), double medial consonant vs. single medial consonant (*annals*:*anal*), and the base form vs. particular suffixed forms (*sane*:*sanity*).

The Bloomfieldian sequencing begins with the /æ/ pronunciation for a, introducing the /e/ pronunciation at a later time with no special emphasis on the relation between /æ/ and /e/ when derived from a. An alternative to this approach is to present both pronunciations at once, working with such pairs as rat:rate, mat:mate, fat:fate, hat:hate and man:mane. Both the associations of a to /æ/ and a to /e/ and the discrimination of the graphemic environments would be emphasized. Whether or not a child first learning to read can handle this task probably depends upon the pedagogy employed. The potential generalization derived from the differentiation approach, however, certainly is greater than that from the simple-sequence method.

There are other areas where we feel new techniques could be introduced, such as in emphasizing the relational unit-marker dichotomy, but we feel that the task ahead in reading research should be primarily in the hands of the psychologist. While several patterns uncovered in our work need further work, we feel that the spelling-to-sound structure presented here is detailed enough to form the basis for a study of human reading habits. The first goal of such research should be to uncover the extent of spelling-to-sound generalization in so called good readers. The next steps would be to study the development of the generalization habits which do exist and to develop methods for teaching the ones that do not exist, but are deemed necessary for good reading.

LIST OF REFERENCES

Bloomfield, Leonard. Language. New York, 1933.

Bolinger, Dwight L. "Visual morphemes," Lg. 22.333-40 (1946).

Bradley, Henry. "On the relations between spoken and written language with special reference to English," British academy 6.211-32 (1914).

Brown, Goold. The grammar of English grammars. 4th ed. New York, 1859.

Ewert, Alfred. The French language. London, 1933.

Francis, W. Nelson. The structure of American English. New York, 1958.

Gibson, E. J., et al. "An analysis of critical features of letters, tested by a confusion matrix," H. Levin, et al. A basic research program on reading. Ithaca, 1963.

Hall, Robert A., Jr. Sound and spelling in English. Philadelphia, 1961.

________. "Graphemics and linguistics," Proceedings of 1962 annual spring meeting of the American Ethnological Society, pp. 53-59.

Haugen, Einar. First grammatical treatise: the earliest Germanic philology. Language monogram No. 25 Baltimore, 1950 .

Hockett, Charles F. A course in modern linguistics. New York, 1958.

Jespersen, Otto H. A modern English grammar on historical principles. 7 vols. Heidelberg, 1909-1931.

Kenyon, John S., and Knott, Thomas. A Pronouncing dictionary of American English. Springfield, Mass., 1951.

Kingdon, Roger. The groundwork of English stress. London, 1958.

Kurath, Hans. A phonology and prosody of Modern English. Ann Arbor, 1964.

Levin, Harry, *et al*. *A basic research program on reading*. Final report, Cooperative research project No. 639. Cornell University, 1963.

________, editor. *Project literacy reports*. Ithaca, 1964, 1965.

McIntosh, Angus. "*The analysis of written Middle English*," *Trans. Phil. Soc.*, 1956, pp. 26-55.

McLaughlin, John C. *A graphemic-phonemic study of a Middle English manuscript*. The Hague, 1963.

Newman, Stanley S. "English suffixation: a descriptive approach," *Word* 4.24-36 (1948)

Pulgram, Ernst. "Phoneme and grapheme: a parallel," *Word* 7.15-20 (1951).

Saussure, Ferdinand de. *Cours de linguistique générale*. Paris, 1916. Translated by Wade Baskin. *Course in general linguistics*. New York, 1959.

Soffietti, James P. "Why children fail to read: a linguistic analysis," *Harvard Edu. Rev*. 25.62-84 (1955).

Stetson, Raymond H. "The phoneme and the grapheme," *Mélanges de linguistique et de philologie offerts à Jacq. van Ginneken*, 353-56 (Paris, 1937).

Thorndike, Edward L. *The teaching of English suffixes*. New York, 1941.

________. *Thorndike century senior dictionary*. New York, 1941.

Uldall, H. J. "Speech and writing," *Acta Linguistica* 4.11-16 (1944).

Vachek, Josef. "Two chapters on written English," *Brno studies in English*, Vol. 1 (Praha, 1959).

________. "On the interplay of external and internal factors in the development of language," *Lingua* 11.433-48 (1962).

Venezky, Richard L. "A computer program for deriving spelling to sound correlations," Masters thesis, Cornell University, 1962. Published in part in Harry Levin, *et al*. *A basic research program on reading*. Ithaca, 1963.

Waldo, George S. "The significance of accentuation in English words," *Proceedings of the ninth international congress of linguists*, 204-10 (The Hague, 1965).

Webster, Noah. An American dictionary of the English language. Revised and enlarged by Chauncey A. Goodrich, Springfield, Mass., 1852.

Webster's new collegiate dictionary. Edited by John P. Bethel. 2d ed. Springfield, Mass., 1956.

Weir, Ruth H. "Formulation of grapheme-phoneme correspondence rules to aid in the teaching of reading," Final report, Cooperative Research project No. S-039 (Stanford, 1964).

________, and Venezky, Richard L. "Rules to aid in the teaching of reading," Final report, Cooperative Research project No. 2584 (Stanford, 1965).

Zachrisson, Robert E. Anglic, a new agreed simplified English spelling. Uppsala, 1930.

________. "Four hundred years of English spelling reform," Studia Neophilologica 4.1-69 (1931)